KNEE RECONSTRUCTION THROUGH EFFICIENT LINEAR PROGRAMMING

M. Sardarescu[1], N. Paragios[1,2], N. Komodakis[1], R. Raymond[4], P. Hernigou[3], and A. Rahmouni[3]

[1] Laboratoire MAS, Ecole Centrale Paris, Châtenay-Malabry, France, [2] Equipe GALEN, INRIA Saclay Ile-de-France, France
[3] Service de Radiologie et d'Imagerie Médicale, [4] Service d'Orthopédie, Hôpital Henri Mondor, Créteil, France

ABSTRACT

In this paper we propose a novel method to recover the 3D shape of the knee and of the prosthesis for the patients that have undergone total knee replacement surgery. We address this using a set of 2D X-Ray images of the knee taken from different viewpoints around it. The problem is then casted to a multiview stereo reconstruction problem and we consider an energy-minimization approach to recover the bone structure. This function is designed to recover a 3D surface which minimizes image and shape-based terms. The cost related to data is defined through the projections of the 3D surface onto 2D image planes while the prior-knowledge is introduced using a 3D model of the knee. The use of Markov Random Fileds as well as recent advances in discrete optimization and in particular the state-of-the-art Fast PD optimization algorithm is considered towards optimal reconstruction. Promising results demonstrate the potentials of our method.

Index Terms— X-ray imaging, multi-view stereo reconstruction, surface evolution, discreet optimization, metric labeling problem

1. INTRODUCTION

For patients who have undergone total knee replacement surgery, both static and dynamic post-operator observations of the knee are of strong interest. In our work we aim to reconstruct the 3D model of the knee (both bone and prosthetics) of a patient from a set of 2D X-ray images. We address this using a 3D multi-view reconstruction approach. The goal is to find the 3D surface which respects the following criteria : *image projection consistency* : its projection onto the image planes should segment both bone and prosthetics from the background and *anatomical constraint* : it must be a valid 3D representation of the shape of the knee joint.

The multi-view stereo reconstruction of the knee from X-ray images has already been addressed in [1]. In such a context one has to address a number of challenges like the presence of occlusions, the non-homogeneity of the bone areas, the lack of texture on the prosthetics areas and the additional noise on the background. All these make impossible the use of classical 3D mutiview segmentation techniques such as visual hull, voxel coloring, space carving or other classical variational techniques. A variational technique based on Geodesic Active Regions has been proposed in [1], but while having good results in the case of prosthetics, when trying also to segment the bone the complexity of the algorithm grows and the results are not the expected ones. Furthermore, the method was quite sensitive to the initial conditions. We tried to overcome this by using prior knowledge on the anatomical shape of the knee.

We model the problem at hand as the minimization of a *discrete energy* defined in the *Metric Labeling Problem* framework (ML problem for short). It contains a term related to the image data (models the image projection consistency constraint), as well as a term related to prior knowledge on the structure of the knee joint (models the anatomical constraint). The state-of-the-art minimization algorithm *Fast PD* [2] is then used for optimizing the objective function and finding the solution.

This article is organized in 4 sections. Section 2 introduces the metric labeling problem. Section 3 presents our solution and gives some results while section 4 provides the conclusions.

2. METRIC LABELING FRAMEWORK

The ML problem was first introduced in [3]. Given a set of objects $\mathbf{P}$, a set of labels $\mathbf{L}$ and a weighted graph $\mathbf{G} = (\mathbf{P}, \mathbf{E}, \mathbf{w})$ defining the relations between the objects, the problem consists in finding a labeling function $f : \mathbf{P} \rightarrow \mathbf{L}$ that minimizes the energy $\mathbf{E}(f)$ (equation 1). The value for assigning a label f_p to an object p is given by $c_p(f_p)$ and expresses the likelihood of associating the label f_p to p. For each edge $(p, q) \in \mathbf{E}$, we associate a separation cost defined by $w_{pq} \cdot d(f_p, f_q)$ reflecting the strength of the relationship between the objects p and q (given by w_{pq}) and the similarity between the labels f_p and f_q (measured by d, which was a distance function in the initial formulation). The intuition behind this cost definition is that one should associate similar labels to strongly related objects.

$$E(f) = \sum_{p \in \mathbf{P}} c_p(f_p) + \sum_{(p,q) \in \mathbf{E}} w_{pq} d(f_p, f_q) \qquad (1)$$

In [2] the metric condition of the d function was relaxed. In fact, for each edge $(p, q) \in \mathbf{E}$ a different function d_{pq}

can be specified, leading to formulation 2 of the energy. In addition, this function d_{pq} must only respect the property : $\forall a, b \in \mathbf{L} \quad d_{pq}(a,b) \geq 0$ and $d_{pq}(a,b) = 0 \Rightarrow a = b$. (In the following sections we shall extend the notion of Metric Labeling problem to this case.)

$$E(f) = \sum_{p \in \mathbf{P}} c_p(f_p) + \sum_{(p,q) \in \mathbf{E}} d_{pq}(f_p, f_q) \qquad (2)$$

Due to the large spectrum of applications which could be modeled using the Metric Labeling approach, over the past decades it has gained significant research attention. However being a NP-hard problem [4], all known algorithms are approximative. To this end a new framework for the optimization of such an energy, called the *Fast PD* (*Fast Primal-Dual*), has been introduced [2]. The framework redefines the ML problem as an *Integer Programming problem*, uses its *Linear Programming* relaxation and takes advantage of the primal-dual properties. It searches for the solution by iteratively generating pairs (x_i, y_i) of solutions for the primal and dual problems ; x_k is the solution of the optimization process if it is a f-approximation of the optimum solution of the primal integral problem (the pair (x_k, y_k) satisfies the *Primal Dual Principle* [2]). The Fast PD algorithm provides both good theoretical guarantees and good practical results.

3. MULTI-VIEW STEREO FROM X-RAYS

We present in this chapter an approach for reconstructing both knee bones and prosthetics for a patient that has undergone total knee replacement surgery from a set of X-ray images. It consists in deforming a 3D surface towards the optimal solution. This can happen using variational formulations [5, 1]. However these methods are sensitive to initial conditions. In order to overcome this limitation we will proceed with a discrete model – we model our problem as a ML problem having as objects a set of control points defining the 3D surface and associating labels to all possible 3D transformations. The basic assumption of our approach is that by deforming the control points in the 3D space, one can recover a solution which satisfies the anatomical constraint while being supported from the data. Such a solution is optimal if the space of deformations is infinite. However, solving labeling problems with infinite number of labels is known to be intractable. We overcome this by using an incremental deformation approach. The same deformations will be considered for all control points. During each iteration we will then seek the optimal solution to the ML problem among these deformations on top of the existing deformation. This corresponds to finding the labeling that minimizes an energy containing both data and shape-prior based costs. Upon convergence, the output of the algorithm is a 3D surface which is sustained by the input image data and respects the anatomical constraints.

3.1. Input

The input X-ray images are taken from a priori known viewpoints around the knee, so we consider that all projections to the images planes are known (no need for calibration). Let $\mathbf{I} = \{I_1, I_2, ..., I_N\}$ be the set of X-ray images, $\mathbf{\Omega} = \{\Omega_1, \Omega_2, ..., \Omega_N\}$ the corresponding 2D domains and $\mathbf{\Pi} = \{\pi_1, \pi_2, ..., \pi_N\}$ the set of projections such that any projection $\pi_i : \mathbf{R}^3 \rightarrow \Omega_i$ transforms a 3D point $(x, y, z) \in \mathbf{R}^3$ to the corresponding pixel $(u_i, v_i) \in \Omega_i$.

We have used a discrete reprezentation of the surface as a 3D triangulated mesh defined by a set of control points (the adjacence of control points does not change during the optimization phase). The advantage of this reprezentation is that each control point has only a local influence on the shape of the surface. Let $\mathbf{P} = \{P_1, P_2, ..., P_n\}$ be the set of control points and $\mathbf{S} = \mathbf{S}(\mathbf{P}) = \mathbf{S}(P_1, P_2, ..., P_n)$ the 3D surface they define.

We achieve the deformation of the 3D surface simply by applying different 3D transformations to each of its control points. To be consistent with the ML problem formulation, we express all these transformations in terms of labels. For a given label w, we can imagine a function f_w that applies a 3D transformation to the coordinates of a control point P (notation $f_w(P)$). As there is no difference between the control points, we have decided to restrict our 3D deformations to translations. We also add to the label set the label 0 corresponding to the identity transformation : $\forall P \in \mathbf{P}, f_0(P) = P$. Note that the number of labels has to be as small as possible in order to have a fast optimization round and as large as possible in order to search in a bigger neighborhood.

Let $\mathbf{L}$ be the set of labels and $w = (w_1, w_2, ..., w_n) \in \mathbf{L}^n$ a vector of n labels (also called a labeling). We denote by f_w the function that applies the transformation f_{w_i} to the control point $P_i : f_w(\mathbf{P}) = \{f_{w_1}(P_1), f_{w_2}(P_2), ..., f_{w_n}(P_n)\}$.

3.2. Energy

The goal of the algorithm is to find the vector of labels that deforms the initial surface to a surface that respects both *image projection consistency* and *anatomical* constraints (see section 1). This can be easily modeled as the minimization of a discrete *energy* (*cost function*) containing two terms : a *data cost* that punishes the surfaces that are non consistent with the input X-ray images and a *shape prior based cost* which punishes the surfaces that are not "smooth" and that are not valid with respect to the anatomical shape of the knee :

$$w^o = \arg\min_{w \in \mathbf{L}} (\alpha \cdot \mathbf{E_{data}} + \beta \cdot \mathbf{E_{prior}}) \qquad (3)$$

To define the data term it is important to notice that each visible control point, which projects to the interior of the object in an image plane, is expected to have coherent regional statistics. On the other hand, each boundary point is expected to have regional discontinuities ($\mathbb{1}_{\Gamma_i}(u, v) = 1$, if the

pixel (u, v) is on the contour of the projected 2D surface $\pi_i(\mathbf{S}(f_w(\mathbf{P})))$, 0 otherwise).

$$\mathbf{E_{data}}(w) = \sum_{\pi_i \in \mathbf{\Pi}} \left\{ \sum_{j=1}^{n} h_i(\pi_i(f_{w_j}(P_j))) \right\} \quad (4)$$

$$h_i(u,v) = \mathbb{1}_{\Gamma_i}(u,v) \cdot g_i(u,v) + (1 - \mathbb{1}_{\Gamma_i}(u,v)) \cdot \gamma \cdot r_i(u,v)$$

In image locations where the boundaries between bone or prosthesis and background exist, one would expect that the distribution of pixel intensities will inherit at least two populations and therefore the entropy is expected to be high (let H be the entropy). Also, the pixels on the boundaries between bone / prosthetics and the background should have high values for the regional cost, while the pixels that are inside should have low values for the same cost (we denote by W_{uv} an image window centered in (u, v)).

$$g_i(u,v) = H_{W_{uv}}(u,v) - \zeta \cdot r_i(I_{(u,v)}) \quad (5)$$

$$H_{W_{uv}}(u,v) = \sum_{(p,q) \in W_{uv}} \mathbf{Pr}[I(pq)] \cdot log(\mathbf{Pr}[I(pq)])$$

For each pixel (u, v) in an image, let $\mathbf{Pr}_{obj}[I(u, v)]$ be the likelihood that the pixel belongs to the object of interest and $\mathbf{Pr}_{back}[I(u, v)]$ that it belongs to the background. The regional cost is smaller in image locations inside the object of interest than on the background (equation 6).

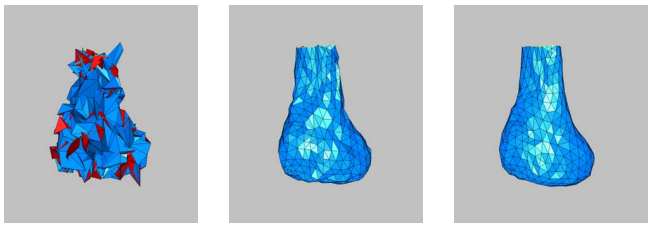
$$r_i(u,v) = \sum_{(p,q) \in W_(u,v)} -\log\left(1 + \frac{\mathbf{Pr}_{obj}[I(p,q)]}{\mathbf{Pr}_{back}[I(p,q)]}\right) \quad (6)$$

To express the shape prior based cost, we begin by defining a graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$, having the control points as nodes ($\mathbf{V} = \mathbf{P}$) and pairs of control points as edges(not only the pairs of adjacent control points on the 3D mesh). Based on some training examples, for each edge $(P_i, P_j) \in \mathbf{E}$ and a positive real value x, we can compute the likelihood that x is the distance between P_i and P_j ; we note this probability $\mathbf{Pr}_{ij}[x]$. Therefore, assuming that either anatomical knowledge is used or actual CT data, we can learn the relative conditional distributions of the position of a control point given the positions of the other control points. The *shape prior-based cost* is defined such that it penalizes the surfaces having high improbable distances between the control points.

$$\mathbf{E_{prior}}(w) = \sum_{(P_i,P_j) \in \mathbf{E}} -\log(\mathbf{Pr}_{ij}[||f_{w_i}(P_i) - f_{w_j}(P_j)||] \quad (7)$$

Equations 3, 4 and 7 lead to the expression 8 of the energy (a typical case of ML energy).

$$\mathbf{E}(w) = \sum_{P_j \in \mathbf{P}} \alpha \sum_{\pi_i \in \mathbf{\Pi}} h_i(\pi_i(f_{w_j}(P_j))) + \sum_{(P_i,P_j) \in \mathbf{E}} -\beta \log\left(\mathbf{Pr}_{ij}\left[||f_{w_i}(P_i) - f_{w_j}(P_j)||\right]\right) \quad (8)$$



(a) Initial surface (b) After 4 iterations (c) Solution surface

Fig. 1. Surface deformation in case of using only shape-prior based cost. The initial surface is an arbitrary 3D triangular mesh. The solution surface (after 9 iterations) is the actual 3D prior model of the knee bone and prosthetics.

We deform the initial surface towards the optimal one using an iterative process. We begin with the initial set of control points $\mathbf{P}^0 = \mathbf{P}$ and the associated surface $\mathbf{S}^0 = \mathbf{S}(\mathbf{P}^0)$. We minimize the energy defined by the equation 8 for the control points $\mathbf{P}^0$ and find the labeling $w^1 \in \mathbf{L}^n$. Then, the control points suffer the 3D transformation defined by the labels w^1 and the result is a new set of control points $\mathbf{P}^1 = f_{w^1}(\mathbf{P}^0)$ that generates a new surface $\mathbf{S}^1 = \mathbf{S}(\mathbf{P}^1)$. The procedure is repeated until convergence – at the final step, we obtain the trivial labeling ($w^t = (0, 0, ..., 0)$) as the result of the optimization algorithm. Even though no theoretical guarantees of convergence are provided, the tests revealed that the solution was found after a small number of iterations.

3.3. Implementation and Results

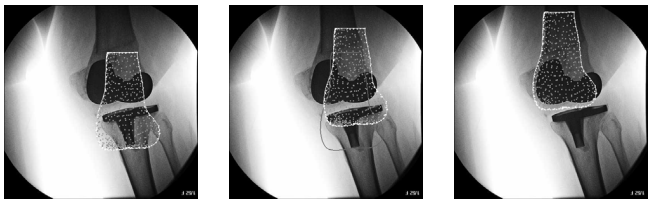
The boundary-based cost (5) and the region cost (6) depend only on the X-ray images (do not depend on the current surface $\mathbf{S}(f_w(\mathbf{P}))$) and their values over each image domain can be pre-computed. Thus, in order to obtain the data cost for P and w, one only needs to project the 3D point $f_w(P)$ to all image domains. We have worked with a surface defined by 500 control points and a set of 130 labels. Even though the time needed for projecting all points to all image domains is important (we have used OpenGL for speeding up the projection process), the overall computational time is not very large due to the fact that the number of iterations needed for convergence is quite small (only 8 in the example below).

Having access to a small number of data sets, we computed the shape-based prior cost using an artificial model of the knee joint. To obtain this model we have relied on the visual hull obtained from a set of input images roughly segmented by hand. This leads to a 3D surface that can be considered as a rough model of the knee joint.

We have first tested our method setting $\alpha = 0$ in equation 8 (only the shape prior based term) and providing as input an arbitrary 3D mesh. The behavior of our algorithm in this case is the one expected, the initial surface suffers a series of 3D transformations and converges rapidly to the prior model (figure 1).

Let's now move to the real problem. We acquire the X-ray images by placing the subject such that his knee is parallel with the Z axis and by rotating the view angle around the knee with 10.885 degrees. The 3D transformations associated to the labels are modified from one iteration to another – this allows a rougher search in the beginning and a finer one at the end (notice however that the label set does not need to change during the iterations).

We begin the algorithm having the shape prior model as the initial surface and an arbitrary initial position. The first steps are done by imposing a high regional and shape prior-based costs (high values for β and γ) while searching in a large neighborhood of the current position (figure 2).



(a) Initial surface (b) After one iteration (c) After 3 iterations

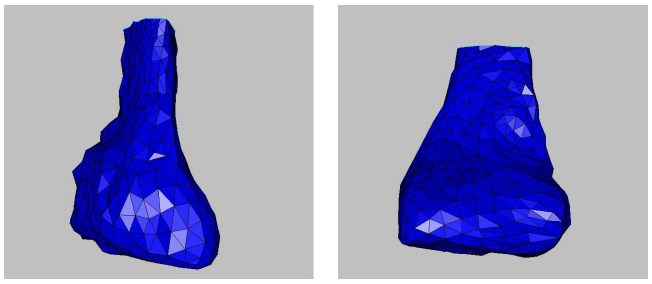
Fig. 2. Surface deformation in case of using strong regional and shape prior based costs. Projections to image with viewangle of 97.965 degrees. The 3D surface suffers translations approaching to the correct pose of the knee. The algorithm converges rapidly.

We give then greater importance to the data term with respect to the shape prior term (high values for α) and we begin a finer search in a smaller neighborhood of the current solution (figures 3 and 4).

(a) After 4 iterations (b) After 6 iterations (c) Solution surface

Fig. 3. Surface deformation using strong data cost. Projections to image with viewangle of 97.965 degrees. The final surface projects correcly to the image data.

Even though the solutions are not ideal, it is important to make some observations. First, the surface deforms such that its image projections are consistent with the image data This validates our data cost formulation. Second, even though the final surface is not smooth enough, one can easily observe that it respects the anatomical geometry of the knee. This shows that our choice of shape prior-based cost is fundamentally correct though it leaves room for further improvement.



(a) (b)

Fig. 4. Final 3D surface. The surface is not as smooth as expected.

4. CONCLUSIONS

In this paper we have addressed the problem of reconstructing of the 3D shape of the knee from a set of X-ray images. Our solution consists in deforming a 3D surface, until its projections to the image planes provide a good separation of the bone and prosthetics from the background in the input X-ray images and such that it is a valid anatomical representation of a human knee. The problem is modeled as a ML problem associating labels to all possible 3D transformations and integrating both *data* and *shape prior-based costs* into the energy. The process of searching for the best surface is iterative, at each iteration seeking the optimal solution among the current deformations on top of the existing deformation. The state-of-the-art optimization algorithm *Fast PD* is used during each optimization round.

The main contribution of our work consists in providing a discrete model for solving the multiview stereo reconstruction problem while making use of a shape prior. Our results show that the concept of prior knowledge improves the reconstruction as opposed to variational approaches .

5. REFERENCES

[1] K. R. Varshney, N. Paragios, A. Kulski, R. Raymond, P. Hernigou, and A. Rahmouni, "Multi-view stereo reconstruction of total knee replacement from x-rays," in *ISBI*. IEEE, 2007, pp. 1148–1151.

[2] N. Komodakis, G. Tziritas, and N. Paragios, "Fast, approximately optimal solutions for single and dynamic mrfs," in *CVPR*. IEEE, 2007, p. in press.

[3] J. M. Kleinberg and E. Tardos, "Approximation algorithms for classification problems with pairwise relationships : Metric labeling and markov random fields," in *FOCS*. IEEE, 1999, pp. 14–23.

[4] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," in *TPAMI*. IEEE, 2001, vol. 23, pp. 1222–1239.

[5] A. Yezzi and S. Soatto, "Stereoscopic segmentation," in *CVPR*. IEEE, 2001, vol. 1, pp. 59–66.